

m. 1 ubet

1644

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/019,348

DATE: 04/14/1999  
TIME: 14:12:09

Input Set: I019348.RAW

This Raw Listing contains the General Information  
Section and up to first 5 pages.

1 <110> APPLICANT: Georgopoulos, Katia  
2 Morgan, Bruce  
3 <120> TITLE OF INVENTION: AIOLOS GENE  
4 <130> FILE REFERENCE: 10287/031001  
5 <140> CURRENT APPLICATION NUMBER: US/09/019,348  
6 <141> CURRENT FILING DATE: 1998-02-05.  
7 <150> EARLIER APPLICATION NUMBER: US 08/733,622  
8 <151> EARLIER FILING DATE: 1996-10-17  
9 <150> EARLIER APPLICATION NUMBER: US 60/017,646  
10 <151> EARLIER FILING DATE: 1996-05-14  
11 <150> EARLIER APPLICATION NUMBER: US 60/005,527  
12 <151> EARLIER FILING DATE: 1995-10-18  
13 <160> NUMBER OF SEQ ID NOS: 29  
14 <170> SOFTWARE: FastSEQ for Windows Version 3.0  
15 <210> SEQ ID NO 1  
16 <211> LENGTH: 1984  
17 <212> TYPE: DNA  
18 <213> ORGANISM: Mus musculus  
19 <220> FEATURE:  
20 <221> NAME/KEY: CDS  
21 <222> LOCATION: (374) ... (1894)  
22 <400> SEQUENCE: 1  
23 caccgagcgca caccgctcgg ctctccttgc gacacgccct catccccggt gtttctcaag 60  
24 tagacgtccc gagacggtcg ctgaggcact gtttccacgc gatcagggtt cctcaggctt 120  
25 gacattcaaa agtgggtgcy gaaccgcgg cactcggagc gtgctttaa gcgcccgcca 180  
26 gccagcgccg ctctaaccct gcgcccggc tgccggcggc tcccgcctg catctgcgc 240  
27 gacgcgaccg agcgatccc gggcctccct gcgccggaa tctcccgcca gccgcgcggg 300  
28 tccccacggc agcagcacgt ggagcggccg cggagcctga gcgacagctg cagcccgcgc 360  
29 ggcccgcggc gac atg gaa gat ata caa ccg act gtg gag ctg aaa agc 409  
30 Met Glu Asp Ile Gln Pro Thr Val Glu Leu Lys Ser  
31 1 5 10  
32 acg gag gag cag cct ctg ccc aca gag agc cca gac gct ctg aat gac 457  
33 Thr Glu Glu Gln Pro Leu Pro Thr Glu Ser Pro Asp Ala Leu Asn Asp  
34 15 20 25  
35 tac agc ttg ccc aaa cct cat gag ata gaa aac gtg gac agt aga gaa 505  
36 Tyr Ser Leu Pro Lys Pro His Glu Ile Glu Asn Val Asp Ser Arg Glu  
37 30 35 40  
38 gcc cca gcc aat gaa gac gaa gat gca gga gaa gat tcg atg aaa gtg 553  
39 Ala Pro Ala Asn Glu Asp Glu Asp Ala Gly Glu Asp Ser Met Lys Val  
40 45 50 55 60  
41 aaa gat gaa tac agc gac aga gat gag aac att atg aag ccg gag ccc 601  
42 Lys Asp Glu Tyr Ser Asp Arg Asp Glu Asn Ile Met Lys Pro Glu Pro  
43 65 70 75  
44 atg gga gat gca gaa gag agt gaa atg cct tac agc tat gca aga gaa 649

ENTERED

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/019,348

DATE: 04/14/1999  
TIME: 14:12:09

Input Set: I019348.RAW

45	Met Gly Asp Ala Glu Glu Ser Glu Met Pro Tyr Ser Tyr Ala Arg Glu	
46	80 85 90	
47	tac agc gac tat gaa agc att aag ctg gag aga cac gtg ccc tat gac	697
48	Tyr Ser Asp Tyr Glu Ser Ile Lys Leu Glu Arg His Val Pro Tyr Asp	
49	95 100 105	
50	aac agc aga cca acc agt ggg aag atg aac tgc gac gtg tgc ggg tta	745
51	Asn Ser Arg Pro Thr Ser Gly Lys Met Asn Cys Asp Val Cys Gly Leu	
52	110 115 120	
53	tcc tgc att agc ttc aac gtc ttg atg gtt cat aag cga agc cat acc	793
54	Ser Cys Ile Ser Phe Asn Val Leu Met Val His Lys Arg Ser His Thr	
55	125 130 135 140	
56	ggc gaa cgc ccg ttc cag tgt aat cag tgc ggg gca tct ttt act cag	841
57	Gly Glu Arg Pro Phe Gln Cys Asn Gln Cys Gly Ala Ser Phe Thr Gln	
58	145 150 155	
59	aaa ggt aac ctc ctc cgt cat att aaa ctg cac acg ggg gaa aaa cct	889
60	Lys Gly Asn Leu Leu Arg His Ile Lys Leu His Thr Gly Glu Lys Pro	
61	160 165 170	
62	ttt aag tgt cac ctc tgc aac tac gca tgc caa agg aga gat gcg ctc	937
63	Phe Lys Cys His Leu Cys Asn Tyr Ala Cys Gln Arg Arg Asp Ala Leu	
64	175 180 185	
65	acg gga cac ctt agg aca cat tct gtg gag aag ccg tac aag tgt gag	985
66	Thr Gly His Leu Arg Thr His Ser Val Glu Lys Pro Tyr Lys Cys Glu	
67	190 195 200	
68	ttc tgc gga aga agc tac aag cag aga agc tcc ctg gag gag cac aag	1033
69	Phe Cys Gly Arg Ser Tyr Lys Gln Arg Ser Ser Leu Glu Glu His Lys	
70	205 210 215 220	
71	gaa cgc tgc cga gct ttt ctt cag aac cct gac ctg ggg gac gct gca	1081
72	Glu Arg Cys Arg Ala Phe Leu Gln Asn Pro Asp Leu Gly Asp Ala Ala	
73	225 230 235	
74	agt gtg gag gca aga cac atc aaa gcc gag atg gga agt gag aga gct	1129
75	Ser Val Glu Ala Arg His Ile Lys Ala Glu Met Gly Ser Glu Arg Ala	
76	240 245 250	
77	ctc gtc ctg gac aga tta gca agc aat gtg gct aag cga aaa agc tcg	1177
78	Leu Val Leu Asp Arg Leu Ala Ser Asn Val Ala Lys Arg Lys Ser Ser	
79	255 260 265	
80	atg cct cag aaa ttc atc ggt gag aag cgg cac tgc ttc gat gcc aac	1225
81	Met Pro Gln Lys Phe Ile Gly Glu Lys Arg His Cys Phe Asp Ala Asn	
82	270 275 280	
83	tac aat ccc ggc tac atg tac gag aag gag aac gag atg atg cag acc	1273
84	Tyr Asn Pro Gly Tyr Met Tyr Glu Lys Glu Asn Glu Met Met Gln Thr	
85	285 290 295 300	
86	cgg atg atg gac caa gcc atc aat aac gcc atc ag : tat cta ggg gct	1321
87	Arg Met Met Asp Gln Ala Ile Asn Asn Ala Ile Ser Tyr Leu Gly Ala	
88	305 310 315	
89	gaa gcc ttc cgc ccc tta gtc cag act ccg cct gct ccc acc tct gag	1369
90	Glu Ala Phe Arg Pro Leu Val Gln Thr Pro Pro Ala Pro Thr Ser Glu	
91	320 325 330	
92	atg gtc cca gtc atc agc agt gtg tac ccc ata gca ctt act cgg gcc	1417
93	Met Val Pro Val Ile Ser Ser Val Tyr Pro Ile Ala Leu Thr Arg Ala	
94	335 340 345	

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/019,348

DATE: 04/14/1999

TIME: 14:12:09

Input Set: I019348.RAW

```

95      gat atg cca atg ggg gcc ccg cag gag atg gaa aag aaa cgg atc ctc      1465
96      Asp Met Pro Met Gly Ala Pro Gln Glu Met Glu Lys Lys Arg Ile Leu
97      350                      355                      360
98      ctg cca gag aag atc ttg cct tct gaa cga ggt ctg tcc ccc aat aac      1513
99      Leu Pro Glu Lys Ile Leu Pro Ser Glu Arg Gly Leu Ser Pro Asn Asn
100     365                      370                      375                      380
101     agt gcc cag gac tcc aca gac acc gac agc aac cac gag gat cgc caa      1561
102     Ser Ala Gln Asp Ser Thr Asp Thr Asp Ser Asn His Glu Asp Arg Gln
103     385                      390                      395
104     cat ctc tac cag caa agc cac gtg gtc ctc ccc cag gcc cgc aat ggg      1609
105     His Leu Tyr Gln Gln Ser His Val Val Leu Pro Gln Ala Arg Asn Gly
106     400                      405                      410
107     atg cct ctt ctg aag gag gtc cct cgc tct ttt gaa ctc ctc aag ccc      1657
108     Met Pro Leu Leu Lys Glu Val Pro Arg Ser Phe Glu Leu Leu Lys Pro
109     415                      420                      425
110     cct ccc atc tgc ctg agg gac tcc atc aaa gtg atc aac aaa gaa ggg      1705
111     Pro Pro Ile Cys Leu Arg Asp Ser Ile Lys Val Ile Asn Lys Glu Gly
112     430                      435                      440
113     gag gtg atg gat gtg ttt cga tgt gac cac tgc cac gtc ctc ttc cta      1753
114     Glu Val Met Asp Val Phe Arg Cys Asp His Cys His Val Leu Phe Leu
115     445                      450                      455                      460
116     gat tat gtg atg ttc acc atc cac atg ggg tgc cat ggt ttc cgt gat      1801
117     Asp Tyr Val Met Phe Thr Ile His Met Gly Cys His Gly Phe Arg Asp
118     465                      470                      475
119     ccc ttt gag tgt aac atg tgt ggc tat cga agc cac gat cgc tat gag      1849
120     Pro Phe Glu Cys Asn Met Cys Gly Tyr Arg Ser His Asp Arg Tyr Glu
121     480                      485                      490
122     ttc tcc tct cac atc gcc aga gga gag cac aga gcc atg ttg aag      1894
123     Phe Ser Ser His Ile Ala Arg Gly Glu His Arg Ala Met Leu Lys
124     495                      500                      505
125     tgagcatctg tctcaatgc gagggtaaac attgtttttt aaagctgatg gtagccttat      1954
126     ccagtagact gaactcaaac ccacctcgag      1984
127     <210> SEQ ID NO 2
128     <211> LENGTH: 507
129     <212> TYPE: PRT
130     <213> ORGANISM: Mus musculus
131     <400> SEQUENCE: 2
132     Met Glu Asp Ile Gln Pro Thr Val Glu Leu Lys Ser Thr Glu Glu Gln
133     1                      5                      10                      15
134     Pro Leu Pro Thr Glu Ser Pro Asp Ala Leu Asn Asp Tyr Ser Leu Pro
135     20                      25                      30
136     Lys Pro His Glu Ile Glu Asn Val Asp Ser Arg Glu Ala Pro Ala Asn
137     35                      40                      45
138     Glu Asp Glu Asp Ala Gly Glu Asp Ser Met Lys Val Lys Asp Glu Tyr
139     50                      55                      60
140     Ser Asp Arg Asp Glu Asn Ile Met Lys Pro Glu Pro Met Gly Asp Ala
141     65                      70                      75                      80
142     Glu Glu Ser Glu Met Pro Tyr Ser Tyr Ala Arg Glu Tyr Ser Asp Tyr
143     85                      90                      95
144     Glu Ser Ile Lys Leu Glu Arg His Val Pro Tyr Asp Asn Ser Arg Pro

```

PAGE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/019,348

DATE: 04/14/1999

TIME: 14:12:09

Input Set: I019348.RAW

145				100				105				110				
146	Thr	Ser	Gly	Lys	Met	Asn	Cys	Asp	Val	Cys	Gly	Leu	Ser	Cys	Ile	Ser
147			115					120				125				
148	Phe	Asn	Val	Leu	Met	Val	His	Lys	Arg	Ser	His	Thr	Gly	Glu	Arg	Pro
149		130					135					140				
150	Phe	Gln	Cys	Asn	Gln	Cys	Gly	Ala	Ser	Phe	Thr	Gln	Lys	Gly	Asn	Leu
151	145					150				155					160	
152	Leu	Arg	His	Ile	Lys	Leu	His	Thr	Gly	Glu	Lys	Pro	Phe	Lys	Cys	His
153				165						170				175		
154	Leu	Cys	Asn	Tyr	Ala	Cys	Gln	Arg	Arg	Asp	Ala	Leu	Thr	Gly	His	Leu
155			180						185			190				
156	Arg	Thr	His	Ser	Val	Glu	Lys	Pro	Tyr	Lys	Cys	Glu	Phe	Cys	Gly	Arg
157		195						200				205				
158	Ser	Tyr	Lys	Gln	Arg	Ser	Ser	Leu	Glu	Glu	His	Lys	Glu	Arg	Cys	Arg
159		210					215					220				
160	Ala	Phe	Leu	Gln	Asn	Pro	Asp	Leu	Gly	Asp	Ala	Ala	Ser	Val	Glu	Ala
161	225					230				235					240	
162	Arg	His	Ile	Lys	Ala	Glu	Met	Gly	Ser	Glu	Arg	Ala	Leu	Val	Leu	Asp
163				245						250				255		
164	Arg	Leu	Ala	Ser	Asn	Val	Ala	Lys	Arg	Lys	Ser	Ser	Met	Pro	Gln	Lys
165			260						265				270			
166	Phe	Ile	Gly	Glu	Lys	Arg	His	Cys	Phe	Asp	Ala	Asn	Tyr	Asn	Pro	Gly
167		275					280					285				
168	Tyr	Met	Tyr	Glu	Lys	Glu	Asn	Glu	Met	Met	Gln	Thr	Arg	Met	Met	Asp
169		290				295					300					
170	Gln	Ala	Ile	Asn	Asn	Ala	Ile	Ser	Tyr	Leu	Gly	Ala	Glu	Ala	Phe	Arg
171	305					310				315					320	
172	Pro	Leu	Val	Gln	Thr	Pro	Pro	Ala	Pro	Thr	Ser	Glu	Met	Val	Pro	Val
173				325						330				335		
174	Ile	Ser	Ser	Val	Tyr	Pro	Ile	Ala	Leu	Thr	Arg	Ala	Asp	Met	Pro	Met
175			340						345				350			
176	Gly	Ala	Pro	Gln	Glu	Met	Glu	Lys	Lys	Arg	Ile	Leu	Leu	Pro	Glu	Lys
177		355					360				365					
178	Ile	Leu	Pro	Ser	Glu	Arg	Gly	Leu	Ser	Pro	Asn	Asn	Ser	Ala	Gln	Asp
179		370				375				380						
180	Ser	Thr	Asp	Thr	Asp	Ser	Asn	His	Glu	Asp	Arg	Gln	His	Leu	Tyr	Gln
181	385					390				395				400		
182	Gln	Ser	His	Val	Val	Leu	Pro	Gln	Ala	Arg	Asn	Gly	Met	Pro	Leu	Leu
183				405						410				415		
184	Lys	Glu	Val	Pro	Arg	Ser	Phe	Glu	Leu	Leu	Lys	Pro	Pro	Pro	Ile	Cys
185			420						425				430			
186	Leu	Arg	Asp	Ser	Ile	Lys	Val	Ile	Asn	Lys	Glu	Gly	Glu	Val	Met	Asp
187		435					440				445					
188	Val	Phe	Arg	Cys	Asp	His	Cys	His	Val	Leu	Phe	Leu	Asp	Tyr	Val	Met
189		450				455				460						
190	Phe	Thr	Ile	His	Met	Gly	Cys	His	Gly	Phe	Arg	Asp	Pro	Phe	Glu	Cys
191	465					470				475				480		
192	Asn	Met	Cys	Gly	Tyr	Arg	Ser	His	Asp	Arg	Tyr	Glu	Phe	Ser	Ser	His
193				485					490				495			
194	Ile	Ala	Arg	Gly	Glu	His	Arg	Ala	Met	Leu	Lys					

PAGE: 5

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/019,348

DATE: 04/14/1999  
TIME: 14:12:09

Input Set: I019348.RAW

```

195                               500                               505
196 <210> SEQ ID NO 3
197 <211> LENGTH: 26
198 <212> TYPE: DNA
199 <213> ORGANISM: Mus musculus
200 <400> SEQUENCE: 3
201      tactaccatc tcacatgggc tgacca                                26
202 <210> SEQ ID NO 4
203 <211> LENGTH: 26
204 <212> TYPE: DNA
205 <213> ORGANISM: Mus musculus
206 <400> SEQUENCE: 4
207      gaccagcaca tggtgacact ctgaaa                                26
208 <210> SEQ ID NO 5
209 <211> LENGTH: 24
210 <212> TYPE: DNA
211 <213> ORGANISM: Mus musculus
212 <400> SEQUENCE: 5
213      gtgtgcgggt tatectgcat tagc                                    24
214 <210> SEQ ID NO 6
215 <211> LENGTH: 24
216 <212> TYPE: DNA
217 <213> ORGANISM: Mus musculus
218 <400> SEQUENCE: 6
219      atcgaagcag tgccgcttct cacc                                    24
220 <210> SEQ ID NO 7
221 <211> LENGTH: 628
222 <212> TYPE: DNA
223 <213> ORGANISM: Homo sapiens
224 <220> FEATURE:
225 <221> NAME/KEY: CDS
226 <222> LOCATION: (1)...(627)
227 <400> SEQUENCE: 7
228      gaa aga gat gag aat gtt tta aag tca gaa ccc atg gga aat gca gaa      48
229      Glu Arg Asp Glu Asn Val Leu Lys Ser Glu Pro Met Gly Asn Ala Glu
230      1          5          10          15
231      gag cct gaa atc cct tac agc tat tca aga gaa tat aat gaa tat gaa      96
232      Glu Pro Glu Ile Pro Tyr Ser Tyr Ser Arg Glu Tyr Asn Glu Tyr Glu
233      20          25          30
234      aac att aag ttg gag aga cat gtt gtc tca ttc gat agt agc agg cca      144
235      Asn Ile Lys Leu Glu Arg His Val Val Ser Phe Asp Ser Ser Arg Pro
236      35          40          45
237      acc agt gga aag atg aac tgc gat gtg tgt gga tta tcc tgc atc agc      192
238      Thr Ser Gly Lys Met Asn Cys Asp Val Cys Gly Leu Ser Cys Ile Ser
239      50          55          60
240      ttc aat gtc tta atg gtt cat aag cga agc cat act ggt gaa cgc cca      240
241      Phe Asn Val Leu Met Val His Lys Arg Ser His Thr Gly Glu Arg Pro
242      65          70          75          80
243      ttc cag tgt aat cag tgt ggg gca tct ttt act cag aaa ggt aac ctc      288
244      Phe Gln Cys Asn Gln Cys Gly Ala Ser Phe Thr Gln Lys Gly Asn Leu

```

Please Note:

Use T n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields for each sequence which presents at least one n or Xaa.



Input Set: I019348.RAW

Line	Error/Warning	Original Text
733	W Invalid/Missing Amino Acid Numbering	
982	W "N" or "Xaa" used: Feature required	Xaa Xaa Ala Ser Asn Val Lys Val Glu Thr G
990	W "N" or "Xaa" used: Feature required	Lys Leu Lys Cys Asp Ile Cys Gly Ile Xaa C
1002	W "N" or "Xaa" used: Feature required	Gln Arg Xaa Ser Leu Glu Glu His Lys Glu A
1004	W "N" or "Xaa" used: Feature required	Glu Ser Met Gly Leu Pro Gly Xaa Xaa Xaa P
1006	W "N" or "Xaa" used: Feature required	Thr Xaa His Xaa Glu Met Ala Glu Asp Leu C
1010	W "N" or "Xaa" used: Feature required	Ser Ser Met Pro Gln Lys Phe Leu Gly Asp L
1012	W "N" or "Xaa" used: Feature required	Pro Tyr Asp Ser Ala Xaa Tyr Glu Lys Glu X
1014	W "N" or "Xaa" used: Feature required	His Val Met Asp Xaa Ala Ile Asn Asn Ala I
1016	W "N" or "Xaa" used: Feature required	Glu Ser Leu Arg Pro Leu Val Gln Thr Pro P
1018	W "N" or "Xaa" used: Feature required	Val Pro Val Ile Ser Pro Met Tyr Gln Leu H
1020	W "N" or "Xaa" used: Feature required	Gly Xaa Pro Arg Ser Asn His Ser Ala Gln A
1022	W "N" or "Xaa" used: Feature required	Leu Leu Leu Leu Ser Lys Ala Lys Xaa Val X
1024	W "N" or "Xaa" used: Feature required	Ser Pro Ser Asn Ser Cys Gln Asp Ser Thr A
1026	W "N" or "Xaa" used: Feature required	Glu Glu Gln Arg Ser Gly Leu Ile Tyr Leu T
1028	W "N" or "Xaa" used: Feature required	Xaa Ala Xaa Xaa Xaa Xaa Xaa Leu Lys Glu G
1030	W "N" or "Xaa" used: Feature required	Xaa Leu Arg Ala Ala Ser Glu Asn Ser Gln A
1032	W "N" or "Xaa" used: Feature required	Ser Thr Ser Gly Glu Gln Xaa Lys Val Tyr L
1034	W "N" or "Xaa" used: Feature required	Val Leu Phe Leu Asp His Val Met Tyr Thr I
1040	W "N" or "Xaa" used: Feature required	His Arg Xaa His Xaa Ser